DOCUMENT RESUME

ED 111 171 FL 005 512

AUTHOR Bordie, John G.
TITLE Direct and Indirect Measures of Language Proficiency.
PUB DATE 1 Jan 73
NOTE 11p.; Paper presented at the meeting of the Linguistic Association of the Southwest (LASSO) (January 1, 1973); Best copy available

EDRS PRICE MF-$0.76 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Elementary Education; Evaluation; *Language Ability; *Language Proficiency; Language Skills; *Language Tests; *Measurement Instruments; *Measurement Techniques; Test Reliability; Verbal Ability; Written Language

ABSTRACT

There is a lack of adequate measurement techniques for testing language proficiency. Researchers compose specific tests for a certain task, but these have only limited general applicability. Often multiple-choice, true-false or fill-in-the-blank tests are used, but these rely heavily on written language and are inadequate for those with poor written but good verbal skills. Such tests tend to compartmentalize learning into components and neglect the overall view, and they generally reflect academic language rather than current vernacular. Norm-referenced, criterion-referenced and non-formal tests involving listing of words following a language cue may be affected by the individual's attitude toward being tested. Indirect tests may solve some of these problems by examining language produced in a non-test situation. Factors studied are sentence length, structural complexity, lexical choice and type-token ration. Occurrences of linguistic mazes and culs-de-sac may also be observed. Second-stage indirect measures such as body language may also be useful. It is possible, however, that proficiency may be a mark of social rather than linguistic status. (CHK)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

BEST COPY AVAILABLE

Direct and In Direct Measures of Language Proficiency

JOHN G. BORDIE - U of Texas at Austin

LASSO 1/1/73

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED111171

Legislation passed by the Texas State Legislature in 1973 calls for the testing of children to determine their proficiency in English prior to their beginning a course of study in the bilingual component of the elementary grades. This legislation specifies proficiency measurement at the entry point, usually kindergarten or first grade, and at any requested exit point up to grade six. Yet the legislature in establishing such a requirement was not aware of the general lack of adequate measurement techniques nor did it indicate developmental factors were of importance at the age being examined. In a previous study of language proficiency measures used in 200 large scale educational experiments, no one measure was used in more than five of the experiments. Most researches chose to use techniques designed for the experiment at hand thereby indicating their disagreement with or their lack of faith in the reliability of existing measures. Whether this signals a lack of trust for all such measures or is symptomatic of uncertainty about factors which are to be measured is not clear but if satisfactory measures are available, researchers do not concede their existence. Reasons for such distrust apparently derive from the feeling expressed by Page not whether the test measures "for what it measures is very well measured, but whether it is measuring the right thing". (F.B. Page, Iowa Test, Revised ed., Buros 6th Mental Measurement Handbook, page 51). The result is the development of many instruments for highly specific purposes having only limited general applicability. With such a lack of comprehensive tests many investigators have chosen to use quick

FL005512

or surface measures in their evaluation of proficiency. Because there are considerations of time or money, many of these devices have been of the paper-pencil variety which evaluate language proficiency on the basis of a two or more choice multiple answer test. These include the True-False question,[1] the multiple answer question,[2] fill in the blank from a preselected list[3] or on the basis of the partial cue, and open ended questions which may be responses to a reading passage,[4] a call for summary or paraphrase, or an essay type question.[5]

These techniques are limited in their heavy emphasis on the written language and are useful only with older and/or literate individuals capable of handling a written stimulus. Such techniques are also limiting for individuals in a culture which stresses oral capability and underemphasizes the ability to write, current teenage America for instance. Where the individual is literate but uses a writing system which is only partially suitable, bridging techniques using a transcriptional system of some variety may be employed - as, for instance, in attempting to measure language progress for an American student in a Japanese language program[6] - such a technique is unsatisfactory for it involves a third writing system and places a premium on the ability to handle the interchange rapidly. Where measures concentrate on oral performance, the method of handling a recording device or one's reaction to being recorded may become problematic.

Oral and written tests of this variety have several defects. The most apparent is the arbitrary segmentation of a unified competence into a series of small, often unrelated components. Such segmentation is justified to the extent that one can measure only one thing at a time. Yet

the fallacy is immediately apparent: when one pronounces, one must pronounce something; when one uses the past tense, one must also control the syntax; when one chooses a specific meaning, this is conveyed through all of the artifacts of language simultaneously. By examining the segments one tends not to see the unity joining the individual items. The most severe outcome of this is the general success of a student in his classwork and his subsequent failure to put the pieces together.

Another limitation is that most proficiency measures tend to reflect academic language usage even when efforts are made to prevent such a happening. It is the exceptional test which reflects and investigates current usage. Slang, fad terms, in group language: all these are notoriously transient. By the time such material appears in a test it is generally outdated and has become formalized: where a test measures lexicon at all it tends to measure mastery of the obsolete or the unusual.

Most such tests of language proficiency are surface measures. That is they rely on language which has been produced in response to a specific cue. Such material is then examined with a view to its standing with respect to an earlier performance by the individual or by some other person or group chosen as a standard. In this sense a surface test can be said to be norm referenced where the degree of similarity or lack of similarity defines the level of accomplishment. Other samples may be compared with some predetermined standard on an acceptable or unacceptable basis. These are criterion referenced. Through appropriate techniques, values for criterion referenced tests can be converted to those for norm referenced tests and vice versa. Whether norm or criterion referenced these

tests are direct because material is elicited for examination. Such tests involve individual awareness of the test in progress causing in many cases, a higher or lower rate of performance than is usual for the individual. Additionally such tests tend, unconsciously or otherwise, to use an unspecified criterion as one referent: the standard or scholastic language of one investigator. Few tests are concerned with, or use, nonstandard language.

Along with the formal direct measures, a variety of non-formal tests have been developed. These non-formal tests generally are non-threatening and may even be regarded as 'fun' by the individual being examined. Non-formal measures are in general as reliable as formal measures though they are regarded with less favor perhaps because they are amusing and do not conform to subconscious attitudes about the serious nature of a test.

An individual may be asked to list as many[8] words as possible, either orally or in writing as an index of lexical availability. He may also be provided with a formal cue which may be graphic[9] - write as many words as you can think of which begin with the letter J - ; phonetic- say as many words as you can which begin with /ʒ/ [10] or semantically categorized - adjectives or color words, etc.[11] The individual may be asked to rearrange something - how many words or phrases can you make out of the word PRESTO without repeating any letter[12] (to date I've found 12 six letter combinations such as POSTER REPOTS etc.). Such techniques inventory lexical awareness.

A particular favorite is a Random word list[13] easily available and with enormous variation. An individual has only to underline all the words

he recognized in a section to indicate immediately his recognition percentage. Or he may be asked to define words in another section for a measure of semantic availability and ability to verbalize. In a random sample, any percentage of the sample is equal to that percentage of the total. Despite their ease of use non-formal measures suffer from the same drawbacks as do formal tests: individual awareness of the test and the tensions incidental to such measurement.

Indirect measures may be used to solve these problems. Such measures are useful when we wish to examine language previously produced in a non-test situation. This provides a more normal sample whose characteristics are not contaminated by extraneous concerns to language.

A number have been of considerable success. Sentence[14] or utterance length correlates very well with actual ability in both oral and written language. Length in written language may not indicate a high level of proficiency in oral language or the reverse but length of sentence wherever found is a significant indicator. Graduate students seem, thus, to be the most proficient. To guard against the inflationary aspects of list sentences,[15] it is preferable to use ~~mean~~ median or mode sentence length rather than the overall average. Closely connected with sentence length is structural complexity[16-17] in which the ratio of complex sentences to simple or compound sentences is an indication of language mastery. Complex here means any sentence with subordination no matter of what variety. A measure related to this of recent popular interest is T(erminal) unit length in which[16-14] the mean length of a main clause plus any associated clauses becomes significant of proficiency as the number increases. The ability to handle a sentence which is structurally complex, indicated by clause

length, is an accepted mark of language proficiency in this society. Precise noncomplex short sentences are also valued. For these sentences, lexical choice is significant with use of words expressing tentativeness, relations between items in a list or a sequence, and conditionality along with a varied vocabulary indicating a greater proficiency. Here the use of "rare" words is deceptive for most words are rare. Francis and Kučera indicate in their study of English that most words have a frequency of occurrence in running text of only 1 1/2 to 2 per million. A more satisfactory category is the type-token ratio based on a sample of 2000-3000 words.[18] In such a passage, the number of different words or types are indicated, then the number of words or tokens are noted. The proportionate ratio established indicates the degree of lexical, and to a certain extent, syntactic mastery.

Another measure is maze frequency.[19] A maze is a verbal tangle which the speaker or writer is unable to resolve with ease and must, as a consequence, use a large number of items to express his views. Economy of effort is a factor here. Related to the maze is the occurence of a cul-de-sac[20] in which the language user, faced by a particular maze abandons the attempt for a start at a different point. The lower the frequency of mazes and cul-de-sacs the more proficient the individual. Mean length of both maze and cul-de-sac also decrease with increasing proficiency. It should be noted that hesitation pauses in oral language do not conform to maze or cul-de-sac frequency. Pause in speech may be (and usually is) an indication of ordering; rehearsing, or search reflecting an individual's preparation for response rather than an inability to respond.

Certain indirect measures, though valuable for other purposes, have little or no correlation with language proficiency: the most obvious is the ability to read and/or write. The so called Mechanical skills, spelling and punctuation[21] in written language, and pronunciation have only a tenuous connection with the ability to use the language. Such areas of proficiency as register or stylistic choice, semantic association, personal delivery, and others are only casually connected to the mechanical skills.

There are second stage indirect measures which are also useful in evaluating proficiency. These measures examine performance in non-language areas for inferential evaluation of language use. The most obvious is body language where movement, attitude, gesture and other similar activities function as counterpoint to actual language use. For written language, the quality of the writer's hand or the choice of perfumed paper are further factors relating to the ability to communicate. Less obvious though equally important are such items as sex - women being generally more proficient than men, - age of associates - through age 30, the older one's associates compared to the individuals age, the more proficient the individual. Other items such as socio-economic status, parental occupation, regional residence, rural/urban distinction, and environmental characteristics are also significant.

Such second stage indirect measures have been attacked from several directions because they describe proficiency from a variable non objective point of view which tends to exclude all items not conforming to expected norms. These criticisms are unnecessarily misleading for there exists no objective description of language proficiency at this time. (For that

matter a definition of language is still unavailable). Even should a definition become available, social factors only slightly understood at present would have to be considered. In closing I would like to quote Leonard Bloomfield[22]. In his article "Literate and Illiterate Speech," (American Speech (1927)2:432-439), referring to his work among the Menomini he states "The nearest approach to an explanation of 'good' and 'bad' language seems to be this, then, that by a cumulation of obvious superiorities both of character and standing, as well as of language, some persons are felt to be better models of conduct and speech than others. Therefore even in matters where the preference is not obvious, the forms which these persons use are felt to have the better flavor. This may be a generally human state of affairs true in every group and applicable to all languages." End quote. Proficiency then may be a mark of social rather than linguistic status.

REFERENCES

Fishman and Cooper. "Alternative Means of Bilingualism," JVLVB VIII: 276-82 (1969).

Fitts and Posner. Human Performance. Brooks Cole, 1969.

Hunt. "Grammatical Structures Written at Three Grade Levels," NCTE (1965).

Loban Studies.

McCarthy, Dorothea. "Language Development in Children," in Manual of Child Psychology, Leonard Carmichael, ed.. New York: John Wiley and Sons, 1954.

Scherer, George A. and Michael Wertheimer. A Psycholinguistic Experiment in Foreign Language Teaching. New York: McGraw-Hill Book Co., 1964.

Valette, Rebecca. Modern Language Testing. New York: Harcourt, Brace and World, 1967.

Yamagiwa. "A Checklist of Tests for Various Types of Proficiency in an Foreign Language," Language Learning VII:3-4 (1956-7).

LASSO HANDOUT
11/2/73

irect and Indirect Measures

1. The plural of MAN is MEN. True False

2. Henry likes to invent excuses for not studying.
Make up Make out Make over Make in

3. _____she own a car? Does, Is, Has

4. According to the paragraph, the proper way to end a question is with rising, falling, level intonation.

5. Describe the pronunciation of words using the TH digraph.

6. Ta˜bako do˜o desu ka--I want a cigarette. How about a cigarette?
He smokes cigarettes.

7. Defines 23 skiddoo, A-OK, Colorless green ideas.

8. Write as many words as you can think of in the next three minutes.

9. Write as many words in the next five minutes as you can that begin with J.

10. Say as many words as you can which begin with / ʒ /.

11. Write as many words or phrases which describe sounds adjectivally.

12. How many different words or phrases can you find in PRESTO without repeating any of the letters?

13. Random Word List from The English Speculum, J.L. Dolby and H.L. Resnikoff, Lockheed Missiles & Space Co., Sunnyvale, California, 1964.

14. Mean written sentence length. ca. 27 words for college seniors.

15. John and Henry and Mary went to Fort Worth to see SCMLA and Lasso and then went to Dallas and saw Six Flags and then went home.

16. The man I had seen opening the door was one of Henry's collegiate friends with whom Jim had been particularly close.

17. I saw the man. He was opening the door. He was Henry's friend.

18. Type/token ratio about 4/10 for average adult.

19. I went to see Henry which he/was had been one of my friends.

20. I went to see Henry which was been mine . . . . .

21. Punctuate: John where Henry had had had had had had

22. L. Bloomfield Literate and Illiterate Speech American Speech 1927 2:432-439